UNIVERSITI TUNKU ABDUL RAHMAN

Department of Mathematics and Actuarial Science

CONTENTS

1 Multiple Random Variable

1.1 Joint Discrete Distributions

In many applications there will be more than one random variable of interest, say $X_1, X_2, \ldots, X_k$. It is convenient mathematically to regard these variables as components of a $k-$dimensional vector, $X = (X_1, X_2, \ldots, X_k)$, which is capable of assuming values $x = (x_1, x_2, \ldots, x_k)$ in a $k-$ dimensional Euclidean space. Note, for example, that an observed value x may be the result of measuring k characteristics once each, or the result of measuring one characteristic k times.

Definition 1. The joint probability density function (joint pdf) of the $k-$dimensional discrete random variable $X = (X_1, X_2, \ldots, X_k)$ is defined to be

$$f(x_1, x_2, \ldots, x_k) = P[X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)$$

for all possible values $x = (x_1, x_2, \ldots, x_k)$ of X.

Theorem 1. A function $f(x_1, x_2, \ldots, x_k)$ is the joint pdf for some vector-valued random variable

$$X = (x_1, x_2, \ldots, x_k)$$

if and only if the following properties are satisfied

1. $f(x_1, x_2, \ldots, x_k) > 0$ for all possible values $x_1, x_2, \ldots, x_k$
2. $\sum_{x_1} \sum_{x_2} \cdots \sum_{x_k} f(x_1, x_2, ..., x_k) = 1$

Definition 2. If the pair (X_1, X_2) of discrete random variables has the joint pdf $f(x_1, x_2)$, then the marginal pdf's of X_1 and X_2 are

$$f_1(x_1) = \sum_{x_2} f(x_1, x_2)$$

and

$$f_2(x_2) = \sum_{x_1} f(x_1, x_2)$$

Example 1.
Let the joint pmf of X_1 and X_2 be defined by

$$p(x_1, x_2) = \frac{x_1 + x_2}{32}, \quad x_1 = 1, 2, x_2 = 1, 2, 3, 4.$$

(a) Display the joint probability distribution of X_1 and X_2 in a table.

(b) Verify that the probability function satisfies Theorem 1.

(c) Find $P(X_1 < X_2)$.

(d) Find $P(X_1 + X_2 = 4)$.

Definition 3. Joint CDF The joint cumulative distribution function of the k random variables $X_1, X_2, \ldots X_k$ is the function defined by

$$F(x_1, x_2, \ldots, x_k) = P[X_1 \leq x_1, \ldots, X_k \leq x_k]$$

Theorem 2. A function $F(x_1, x_2)$ is a bivariate CDF if and only if

- $\lim_{x_1 \to -\infty} F(x_1, x_2) = F(-\infty, x_2) = 0 \; \forall \; x_2$
- $\lim_{x_2 \to -\infty} F(x_1, x_2) = F(x_1, -\infty) = 0 \; \forall \; x_1$
- $\lim_{x_1 \to \infty, x_2 \to \infty} F(x_1, x_2) = F(\infty, \infty) = 1 \; \forall \; x_1, x_2$
- $F(b, d) - F(b, c) - F(a, d) + F(a, c) \geq 0 \; \forall \; a < b, c < d$
- $\lim_{h \to 0^+} F(x_1 + h, x_2) = \lim_{h \to 0^+} F(x_1, x_2 + h) = F(x_1, x_2)$

Example 2.
A local supermarket has three checkout counters. Two customers arrive at the counters at different times when the counters are serving no other customers. Each customer chooses a counter at random, independently of the other. Let X_1 denote the number of customers who choose counter 1 and X_2, the number who select counter 2.

(a) Find the joint probability function of X_1 and X_2.

(b) Find $F(-1, 2)$, $F(1.5, 2)$, and $F(5, 7)$.

Example 3. If X and Y are discrete random variables with joint pdf

$$f(x,y) = c\frac{2^{x+y}}{x!y!} \quad x = 0,1,2,\ldots; y = 0,1,2,\ldots$$

and zero otherwise.

(a) Find the constant c.

(b) Find the marginal pdf's of X and Y.

1.2 Joint Continuous Distributions

Definition 4. A $k-$dimensional vector valued random variable $X = (X_1, X_2, \ldots, X_k)$ is said to be continuous if there is a function $f(x_1, x_2, \ldots, x_k)$, called the joint probability density function (joint pdf), of X, such that the joint CDF can be written as

$$F(x_1, x_2, \ldots, x_k) = \int_{-\infty}^{x_k} \cdots \int_{-\infty}^{x_1} dt_1 \cdots dt_k$$

$\forall\ x = (x_1, x_2, \ldots, x_k)$.

Theorem 3. Any function $f(x_1, x_2, \ldots, x_k)$ is a joint pdf of a $k-$dimensional random variable if and only if

1. $f(x_1, x_2, \ldots, x_k) \geq 0\ \forall\ x_1, x_2, , \ldots, x_k$
2. $\int_{\infty}^{\infty} \cdots \int_{\infty}^{\infty} f(x_1, x_2, \ldots, x_k) dx_1 \cdots dx_k = 1$

Example 4.
Let X_1 denote the concentration of a certain substance in one trial of an experiment, and X_2 the concentration of the substance in a second trial of the experiment. Assume that the joint pdf is given by $f(x_1, x_2) = 4x_1x_2; 0 < x_1 < 1, 0 < x_2 < 1$, and zero otherwise.

(a) Find the joint CDF.

(b) Find $P\left[\frac{X_1+X_2}{2} < 0.5\right]$.

Definition 5. If the pair (X_1, X_2) of continuous random variables has the Joint pdf $f(x_1, x_2)$, then the marginal pdf's of X_1 and X_2 are

$$f_1(x_1) = \int_0^\infty f(x_1, x_2)dx_2$$

and

$$f_2(x_2) = \int_0^\infty f(x_1, x_2)dx_1$$

Definition 6. 1f $X = (X_1, X_2, \ldots, X_k)$ is a $k-$dimensional random variable with joint CDF $F(x_1, x_2, \ldots, x_k)$, then the marginal CDF of X is

$$F_j(x_j) = \lim_{x_i \to \infty, \text{all } i \neq j} F(x_1, \ldots, x_j, \ldots, x_k)$$

Furthermore, if X is discrete, the marginal pdf is

$f_j(x_j) = \sum \cdots_{\text{all } i \neq j} \sum f(x_1, \ldots, x_j, \ldots, x_k)$

and if X is continuous, the marginal pdf is

$$f_j(x_j) = \int \cdots_{\text{all } i \neq j} \int f(x_1, \ldots, x_j, \ldots, x_k)dx_1 \ldots dx_k$$

Example 5.
Suppose that a radioactive particle is randomly located in a square with sides of unit length. That is, if two regions within the unit square and of equal area are considered, the particle is equally likely to be in either region. Let X_1 and X_2 denote the coordinates of the particle's location. A reasonable model for the relative frequency histogram for X_1 and X_2 is the bivariate analogue of the univariate uniform density function:

$$f(x_1, x_2) = \begin{cases} 1, & 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

(a) Sketch the probability density surface.

(b) Find $F(.2, .4)$.

(c) Find $P(.1 \leq X_1 \leq .3, 0 \leq X_2 \leq .5)$

Example 6. The joint probability density function of X_1 and X_2 is

$$f(x_1, x_2) = \begin{cases} 3x_1, & 0 \leq x_2 \leq x_1 \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

(a) Sketch the probability density surface.

(b) Find $P(0 \leq X_1 \leq .5, X_2 \geq 0.25)$.

1.3 Conditional Distributions

Definition 7. Conditional pdf If X_1 and X_2 are discrete or continuous random variables with joint pdf $f(x_1, x_2)$, then the conditional probability density function (conditional pdf) of X_2 given $X_1 = x_1$ is defined to be

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$

for values x_1 such that $f_1(x_1) > 0$ and zero otherwise.

Similarly, the conditional pdf of X_1 given $X_2 = x_2$ is defined to be

$$f(x_1|x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$

for values x_2 such that $f_2(x_2) > 0$ and zero otherwise.

Theorem 4. If X_1 and X_2 are random variables with joint pdf $f(x_1, x_2)$ and marginal pdf's $f_1(x_1)$ and $f_2(x_2)$, then

$$f(x_1, x_2) = f_1(x_1)f(x_2|x_1) = f_2(x_2)f(x_1|x_2)$$

and if X_1 and X_2 are independent, then

$$f(x_2|x_1) = f_2(x_2)$$

and

$$f(x_1|x_2) = f_1(x_1)$$

Example 7.
Let
$$f(x_1,x_2,x_3,x_4)=\begin{cases}\frac{3}{4}(x_1^2+x_2^2+x_3^2+x_4^2), & 0<x_i<1, i=1,2,3,4\\ 0, & \text{otherwise}\end{cases}$$

(a) Find the marginal pdf of (X_1, X_2).

(b) Find the conditional pdf of (X_3, X_4) given $X_1 = \frac{1}{3}$ and $X_2 = \frac{2}{3}$.

Example 8. The joint density function of X_1 and X_2 is given by
$$f(x_1,x_2)=\begin{cases}30x_1x_2^2, & x_1-1\le x_2\le 1-x_1, 0\le x_1\le 1\\ 0, & \text{otherwise}\end{cases}$$

(a) Show that the marginal density of X_1 is a beta density with $a = 2$ and $b = 4$.

(b) Derive the marginal density of X_2.

(c) Derive the conditional density of X_2 given $X_1 = x_1$.

(d) Find $P(X_2 > 0|X_1 = .75)$.

1.4 Independent Random Variables

Definition 8. Independent Random Variables Random variables $X_1, \ldots, X_k$ are said to be independent if for every $a_i < bi$,

$$P(a_1 \leq x_1 \leq b_1, \cdots, a_k \leq x_k \leq b_k) \\ = \prod_{i=1}^{k} P(a_i \leq x_i \leq b_i)$$

Theorem 5. Random variables $X_1, \ldots, X_k$ are independent if and only if the following properties holds:

$$F(x_1, \ldots, x_k) = F_1(x_1) \cdots F_k(x_k)$$

$$f(x_1, \ldots, x_k) = f_1(x_1) \cdots f_k(x_k)$$

where $F_i(x_i)$ and $f_i(x_i)$ are the marginal CDF and pdf of X, respectively,

Theorem 6. Two random variables X_1 and X_2 with joint pdf $f(x_1, x_2)$ are independent if and only if:

1. The "support set" $\{(x_1, x_2) | f(x_1, x_2) > 0]$, is a Cartesian product, $A \times B$, and
2. The joint pdf can be factored into the product of functions of x_1 and x_2, $f(x_1, x_2) = g(x_1)h(x_2)$

Example 9. The joint pdf of a pair X_1 and X_2 is

$$f(x_1, x_2) = 8x_1x_2, 0 < x_1 < x_2 < 1$$

and zero otherwise. Are X_1 and X_2 independent?

Example 10. Consider now a pair X_1 and X_2 with joint pdf

$$f(x_1, x_2) = x_1 + x_2, 0 < x_1 < 1, 0 < x_2 < 1$$

and zero otherwise. Are X_1 and X_2 independent?

Example 11. The joint distribution of X_1 and X_2 is given by the entries in the following table.

	x_2	
x_1	0	1
0	0.12	0.28
1	0.18	0.42

Show that X_1 and X_2 are independent.

Example 12.
The joint distribution of X_1 and X_2 is given by the entries in the following table.

	x_2		
x_1	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

Is X_1 independent of X_2?

Example 13. Let

$$f(x_1, x_2) = \begin{cases} 2, & 0 \leq x_2 \leq x_1 \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Show that X_1 and X_2 are dependent.

1.5 The Expected Value of a Function of Random Variables

Definition 9. If $X = (X_1, \ldots, X_k)$ has a joint pdf $f(x_1, \ldots, x_k)$, and if $Y = u(X_1, \ldots, X_k)$ is a function of X, then $E(Y) = E[u(X_1, \ldots, X_k)]$, where

$$E_X[u(X_1, \ldots, X_k)] = \sum_{x_1} \cdots \sum_{x_k} u(x_1, \ldots, x_k) f(x_1, \ldots, x_k)$$

if X is discrete, and

$$E_X[u(X_1, \ldots, X_k)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, \ldots, x_k) f(x_1, \ldots, x_k)$$

if X is continuous.

Theorem 7. If X_1 and X_2 are random variables with joint pdf $f(x_1, x_2)$, then

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

It is possible to combine the preceding theorems to show that if $a_1, a_2, \ldots, a_k$, are constants and $X_1, X_2, \ldots, X_k$ are jointly distributed random variables, then

$$E\left(\sum_{i=1}^{k} a_i X_i\right) = \sum_{i=1}^{k} E(X_i)$$

Theorem 8. If X and Y are independent random variables and $g(x)$ and $h(y)$ are functions, then

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

It is possible to generalize this theorem to more than two variables, Specifically, if $X_1, X_2, \ldots, X_k$ are independent random variables, and $u_1(x_1), \ldots, u_k(x_k)$ are functions, then

$$E[u_1(X_1) \cdots U_k(X_k)] = E[u_1(X_1)] \cdots E[u_k(X_k)]$$

Example 14.

The joint distribution of X_1 and X_2 is given by the entries in the following table.

	x_2		
x_1	0	1	2
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$
1	$\frac{2}{9}$	$\frac{2}{9}$	0
2	$\frac{1}{9}$	0	0

(a) Find $E(X_1)$

(b) Find $V(X_1)$

(c) Find $E(X_1X_2)$

Example 15.

Let $f(x_1, x_2) = \begin{cases} 2x_1, & 0 \le x_1 \le 1; 0 \le x_2 \le 1 \\ 0, & \text{otherwise} \end{cases}$

(a) Find $E(X_1)$

(b) Find $V(X_1)$

(c) Find $E(X_1^2X_2)$

Example 16.

Let $f(x_1, x_2) = \begin{cases} 2x_1, & 0 \le x_1 \le 1; 0 \le x_2 \le 1 \\ 0, & \text{otherwise} \end{cases}$

(a) Find $E(X_2)$

(b) Find $E(X_2 - X_1)$

(c) Find $E(5X_1 + 6X_2 - 2X_1^2X_2)$

Example 17.
Suppose X_1 and X_2 are independent random variables, $E(X_1) = 2$ and $E(X_2) = \frac{1}{3}$. Find $E(X_1X_2)$.

Definition 10. The covariance of a pair of random variables X and Y is defined by

$$Cov(X, Y) = E[(X - \mu_X))(Y - \mu_Y)]$$

Another common notation for covariance is σ_{XY}.

Theorem 9. If X and Y are random variables and a and b are constants, then

- $Cov(aX, bY) = abCov(X, Y)$
- $Cov(X + a, Y + b) = Cov(X, Y)$
- $Cov(X, aX + b) = aV(X)$

Theorem 10. If X and Y are random variables, then

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

and $Cov(X, Y) = 0$ whenever X and Y are independent.

Definition 11. If X and Y are random variables with variances σ_X^2 and σ_Y^2 and covariance $\sigma_{XY} = Cov(X, Y)$, then the conelation coefficient of X and Y is

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

The random variables X and Y are said to be uncorrelated if $\rho = 0$; otherwise they are said to be correlated.

Theorem 11. If ρ is the correlation coefficient of X and Y, then

$$-1 \leq \rho \leq 1$$

and $\rho = \pm 1$ if and only if $Y = aX + b$ with probability 1 for some $a \neq 0$ and b.

Theorem 12. If X_1 and X_2 are random variables with joint pdf $f(x_1, x_2)$, then

$$V(X_1+X_2) = V(X_1)+V(X_2)+2Cov(X_1, X_2)$$

and

$$V(X_1 + X_2) = V(X_1) + V(X_2)$$

whenever X_1 and X_2 are independent.

It also can be verified that if $a_1, a_2, \ldots, a_k$, are constants and $X_1, X_2, \ldots, X_k$, are random variables, then

$$V(\sum_{i=1}^{k} a_i X_i) = \sum_{i=1}^{k} a_i^2 V(X_i) + 2 \sum \sum_{i<k} a_i a_j Cov(X_i, X_j)$$

and if $X_1, X_2, \ldots, X_k$ are independent, then

$$V(\sum_{i=1}^{k} a_i X_i) = \sum_{i=1}^{k} a_i^2 V(X_i)$$

Example 18. X_1 and X_2 have joint density given by

$$f(x_1, x_2) = \begin{cases} 2x_1, & 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Find $Cov(X_1, X_2)$.

Example 19.

Let $f(x, y) = 6x, 0 < x < y < 1$, and zero otherwise. Find $Cov(X, Y)$.

Example 20.

Let X and Y be discrete random variables with joint pdf $f(x, y) = \frac{4}{5xy}$ if $x = 1, 2$ and $y = 2, 3$, and zero otherwise. Find $Cov(X, Y)$.

Example 21. Let X_1 and X_2 be discrete random variables with joint probability distribution as show in table below. Show that X_1 and X_2 are dependent but have zero covariance.

	x_2		
x_1	-1	0	1
-1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$
0	$\frac{3}{16}$	0	$\frac{3}{16}$
1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$

1.6 Conditional Expectation

Definition 12. If X and Y are jointly distributed random variables, then the conditional expectation of Y given $X = x$ is given by

$$E(Y|x) = \sum_y yf(y|x) \text{ if } X \text{ and } Y \text{ are discrete}$$

$$E(Y|x) = \int yf(y|x)dy \text{ if } X \text{ and } Y \text{ are continuous}$$

Example 22. Below is a table giving a joint probability function for discrete random variables X_1 and X_2.

x_1 \ x_2	3	4	5	6
4	.1	.05	.05	0
3	.05	0.2	0.2	0
2	0	0	.2	.05
1	0	0	0	.1

(a) Find the conditional mean of X_2 given $X_1 = 4$, $E[X_2|X_1 = 4]$.

(b) Find the conditional variance of X_2 given $X_1 = 4$, $V[X_2|X_1 = 4]$.

Example 23. Let X_1 and X_2 have the joint pdf

$$f(x_1, x_2) = \begin{cases} 1, & 0 < x_2 < 2x_1, 0 < x_1 < 1 \\ 0, & \text{otherwise} \end{cases}$$

Find $E(X_2|X_1 = x_1)$.

Theorem 13. If X and Y are independent random variables, then $E(Y|x) = E(Y)$ and $E(X|y) = E(X)$.

Theorem 14.
Let X and Y denote random variables. Then

$$\boxed{E(X) = E[E(X|Y)]}$$

where, on the right hand side, the inside expectation is with respect to the conditional distribution of X given Y, and the outside expectation is with respect to the distribution of Y.

Theorem 15.
Let X and Y denote random variables and $h(x, y)$ is a function. Then

$$E[h(X,Y)] = E_Y[E(h(X,Y)|Y)]$$

or

$$E[h(X,Y)] = E_X[E(h(X,Y)|X)]$$

Definition 13. The conditional variance of Y given $X = x$ is given by

$$V(Y|x) = E\{[Y - E(Ylx)]2|x\}$$

An equivalent form,is

$$V(Y|x) = E(Y^2|x) - [E(Y|x)]^2$$

Theorem 16.
Let X and Y denote random variables. Then

$$\boxed{V(X) = E[V(X|Y)] + V[E(X|Y)]}$$

Example 24. A quality control plan for an assembly line involves sampling $n = 10$ finished items per day and counting X, the number of defectives. If p denotes the probability of observing a defective, then X has a binomial distribution, assuming that a large number of items are produced by the line. But p varies from day to day and is assumed to have a uniform distribution on the interval from 0 to $\frac{1}{4}$. Find the expected value and variance of X.

Example 25. If $X_2|X_1 = x_1 \sim POI(x_1)$, and $X_1 \sim EXP(1)$, find $E(X_2)$ and $V(X_2)$.

Example 26. Let X_1 be the number of customers arriving in a given minute at the drive-up window of a local bank, and let X_2 be the number who make the withdrawals. Assume X_1 is Poisson distributed with expected value $E(X_1) = 3$, and that the conditional expectation and variance X_2 given $X_1 = x_1$ are $E(X_2|x_1) = \frac{x_1}{2}$ and $V(X_2|x_1) = \frac{x_1+1}{3}$. Find

(a) $E(X_2)$

(b) $V(X_2)$

(c) $E(X_1X_2)$

1.7 Extended Hypergeometric Distribution

Suppose that a collection consists of a finite number of items N and that there are $k+i$ different types; M_1 of type 1, M_2 of type 2, and so on. Select n items at random without replacement, and let X_i be the number of items of type i that are selected. The vector $X = (X_1, X_2, \ldots, X_k)$ has an extended hypergeometric distribution and a joint pdf of the form

$$f(x_1, x_2, \ldots, x_k) = \frac{\binom{M_1}{x_1}\binom{M_2}{x_2}\cdots\binom{M_{k-1}}{x_{k-1}}\binom{M_k}{x_k}}{\binom{N}{n}}$$

for all $0 \leq x \leq M_1$, where $M_{k+1} = N - \sum_{i=1}^{k} M_i$ and $x_{k+1} = n$. A special notation for this is

$$X \sim HYP(n, M_1, M_2, \ldots, M_k, N)$$

Example 27. A bin contained 1000 flower seeds and 400 were red flowering seeds. Of the remaining seeds, 400 are white flowering and 200 are pink flowering. If 10 seeds are selected at random without replacement, then the number of red flowering seeds, X_1, and the number of white flowering seeds, X_2, in the sample are jointly distributed discrete random variables.

(a) Find the joint pdf of the pair (X_1, X_2).

(b) Find the probability of obtaining exactly two red, five white, and three pink flowering seeds.

1.8 Multinomial Distribution

Suppose that there are $k+1$ mutually exclusive and exhaustive events, say $E_1, E_2, \ldots, E_k, E_{k+l}$, which can occur on any trial of an experiment, and let $p_i = P(E_i)$ for $i = 1, 2, \ldots, k+1$. On n independent trials of the experiment, we let X_i be the number of occurrences of the event E_i. The vector $X = (X_1, X_2, \ldots, X_k)$ is said to have the multinomial distribution which has a joint pdf of the form

$$f(x_1, x_2, \ldots, x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

Theorem 17. If $X = (X_1, X_2, \ldots, X_k)$ have a multinomial distribution with parameters n and $p_1, p_2, \ldots, p_k$, then

1. $E(X_i) = np_i$, $V(X_i) = np_i q_i$
2. $Cov(X_s, X_t) = -np_s p_t$, if $s \neq t$

Example 28. According to recent census figures, the proportions of adults (persons over 18 years of age) in the United States associated with five age categories are as given in the following table.

Age	Proportion
18-24	.18
25-34	.23
35-44	.16
45-64	.27
65 & above	.16

If these figures are accurate and five adults are randomly sampled, find the probability that the sample contains one person between the ages of 18 and 24, two between the ages of 25 and 34, and two between the ages of 45 and 64.

Example 29. A large lot of manufactured items contains 10% with exactly one defect, 5% with more than one defect, and the remainder with no defects. Ten items are randomly selected from this lot for sale. If X_1 denotes the number of items with one defect and X_2, the number with more than one defect, the repair costs are $X_1 + 3X_2$. Find the mean and variance of the repair costs.

1.9 Bivariate Normal Distribution

A pair of continuous random variables X and Y is said to have a bivariate normal distribution if it has a joint pdf of the form

$$
\begin{aligned}
f(x,y) = {} & \tfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\
& \times \exp\{-\tfrac{1}{2(1-\rho^2)}[(\tfrac{x-\mu_1}{\sigma_1})^2 + (\tfrac{y-\mu_2}{\sigma_2})^2 \\
& -2\rho(\tfrac{x-\mu_1}{\sigma_1})(\tfrac{y-\mu_2}{\sigma_2})]\}, x \in R, y \in R
\end{aligned}
$$

A special notation for this is

$$(X,Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$$

which depends on five parameters, $\mu_1, \mu_2 \in R$, $\sigma_1^2 > 0, \sigma_2^2 > 0$ and $-1 < \rho < 1$.

Theorem 18. If $(X,Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then $x_1 \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$.

Theorem 19. If $(X,Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then

1. conditional on $X = x$,

$$Y|x \sim N(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2))$$

2. conditional on $Y = y$,

$$X|y \sim N(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2), \sigma_1^2(1-\rho^2))$$

Example 30. Let X_1 and X_2 be independent normal random variables, $X_i \sim N(\mu_i, \sigma_i^2)$, and let $Y_1 = X_1$ and $Y_2 = X_1 + X_2$.

(a) What are the means, variances, and correlation coefficient of Y_1 and Y_2?

(b) Find the conditional distribution of Y_2 given $Y_1 = y_1$.

1.10 Joint Moment Generating Function

The joint MGF of $X = (X_1, \ldots, X_k)$, if it exists, is defined to be

$$M_X(t) = E\left[\exp\left(\sum_{i=1}^{k} t_i X_i\right)\right]$$

Note that it also is possible to obtain the MGF of the marginal distributions from the joint MGF. For example,

$$M_X(t_1) = M_{X,Y}(t_1, 0)$$

$$M_Y(t_2) = M_{X,Y}(0, t_2)$$

Theorem 20. If $M_{XY}(t_1, t_2)$ exists, then the random variables X and Y are independent if and only if

$$M_{XY}(t_1, t_2) = M_X(t_1) M_Y(t_2)$$

Example 31. Suppose that X and Y are continuous with joint pdf $f(x, y) = 2e^{-x-y}$ if $0 < x < y < \infty$ and zero otherwise.

(a) Derive the joint MGF of X and Y.

(b) Derive the MGF of X and Y respectively.